

Crossing Statistic: Bayesian interpretation, model selection and resolving dark energy parametrization problem

Arman Shafieloo

*Institute for the Early Universe, Ewha Womans University
Seoul, 120-750, South Korea
E-mail: arman@ewha.ac.kr*

ABSTRACT: By introducing Crossing functions and hyper-parameters I show that the Bayesian interpretation of the Crossing Statistics [1] can be used trivially for the purpose of model selection among cosmological models. In this approach to falsify a cosmological model there is no need to compare it with other models or assume any particular form of parametrization for the cosmological quantities like luminosity distance, Hubble parameter or equation of state of dark energy. Instead, hyper-parameters of Crossing functions perform as discriminators between correct and wrong models. Using this approach one can falsify any assumed cosmological model without putting priors on the underlying actual model of the universe and its parameters, hence the issue of dark energy parametrization is resolved. It will be also shown that the sensitivity of the method to the intrinsic dispersion of the data is small that is another important characteristic of the method in testing cosmological models dealing with data with high uncertainties.

KEYWORDS: [Supernovae](#), [dark energy](#), [cosmological parameter estimation](#).

Contents

1. Introduction	1
2. Method and Analysis	2
3. Conclusion	7

1. Introduction

One of the main goals of physical cosmology is to reconstruct the expansion history of the universe and finding the actual model of dark energy. There have been many approaches introduced in the cosmological community to reconstruct the expansion history of the universe and the properties of dark energy. One can categorize them in general to parametric and non-parametric methods [2]. Non-parametric methods are least bias by initial assumptions but in many cases they are not efficient and estimation of the errors can be also problematic [3–16]. Contrary to non-parametric methods, parametric methods are easy to use and they may seem so efficient in putting tight constraints on the cosmological quantities but if the assumed parameterization being wrong, which can be easily the case since the actual model of the universe is not yet known, then they can be completely misleading and result to wrong conclusions. See [17–21] for details of data analysis and methods of parametric reconstruction of the properties of dark energy using supernovae data.

In this paper I introduce Crossing functions and hyper-parameters, motivated by the idea of the Crossing Statistic [1], as a tool to discriminate between correct and wrong models. It is known that if an assumed model is indeed the correct one to describe Gaussian distributed data, then the histogram of the normalized residuals should also have a Gaussian distribution, with zero mean and a standard deviation of 1. If the histogram instead exhibits significant deviation from the a Gaussian distribution, however, then this can be used to rule out the assumed model. The Crossing Statistic [1] pushes this well known idea from statistical analysis a step further by pointing to the fact that not only should the whole sample of residuals have a Gaussian distribution around the mean, but so should any continuous sub-sample. In other words we have shown that catching the systematic correlations in

the residuals can be used to falsify assumed models. In the current paper I introduce the Bayesian interpretation of the Crossing Statistic looking for the systematic correlations in the residuals by comparing an assumed model with its own possible variations. In this approach observable quantities of any assumed model is multiplied by a Crossing function and then is fitted to the data and the resultant values of the hyper-parameters of the Crossing function decide about consistency of the model and the data. In fact derived values of the Crossing hyper-parameters indicate that how smooth deviations from the assumed model can improve the fit to the data. If improvement in the fit to the data being significant by smooth deviation from the assumed model, it hints towards the inconsistency of the model and the data. All these can be done in the framework of χ^2 statistics to define the confidence limits. Using supernovae type Ia simulated data I will show how efficiently Crossing Statistics can discriminate between cosmological models while by using the standard statistics we need to assume a parametric form for the cosmological quantities. The core idea behind this work is to suggest that instead of parametrizing cosmological quantities like luminosity distance, Hubble parameter or equation of state of dark energy, we can simply falsify any assumed cosmological model by parameterizing the possible deviations from its actual predictions for the cosmological observables.

In the following I will introduce the Crossing functions and their hyper-parameters and argue how they can be used as discriminators between cosmological models. I will then use simulated data to show how effectively Crossing Statistics can distinguish among cosmological models. Minimal sensitivity of the method to the intrinsic dispersion of the data will be discussed later before I summarize the results at the end.

2. Method and Analysis

If the real Universe differs from the assumed theoretical model, one would hope that it would be possible to develop a statistical test that would be able to pick up on this. To these ends we consider the ‘crossings’ between the predictions of a given model, and the real Universe from which the data has been derived. The idea of Crossing Statistics is based on the fact that any wrong assumed model and the data cross each other at one or two or maybe even more points within the range of data. Here I argue that since within the frame work of FLRW universe the cosmic distances (distance modules) monotonically increases by redshift, hence any two cosmological models can become indistinguishable (up to a very high precision of the data) if we multiply the distance modules of one of them by a function form with a certain degree and for some particular coefficients. The Bayesian interpretation of Crossing Statistic is in fact hidden in these two prior assumptions that 1) $\mu(z)$ increases by

redshift monotonically for all cosmological models hence there are no high frequency fluctuations in $\mu(z)$ and 2) since the distance modules of all cosmological models increases by redshift, $\mu(z)$ of any two cosmological models can become so close to each other at all redshifts up to an indistinguishable level if we multiply $\mu(z)$ of one of them to a suitable smooth function of degree n with some particular values for the coefficients. Coefficients of these functions are in fact the Crossing hyper-parameters that if all of them be simultaneously consistent to zero value then we consider the assumed model to be consistent to the data.

Since the behavior of matter density and its effect on the expansion history of the universe is known, for any assumed model of dark energy we keep Ω_{0m} as a free parameter along with Crossing hyper-parameters of the function form. We multiply the luminosity distance given by our assumed dark energy model $\mu_{DE}(\Omega_{0m}, z)$ to the function $T_N(C_1, \dots, C_N, z)$ and fit this new function to the data using χ^2 minimization and find the best fit point in the hyper-parameter space and derive the confidence limits.

Crossing function is a parametrization proposed to pick up the possible deviations from the assumed theoretical model using the data. The basic difference between the Bayesian interpretation of the Crossing statistic and usual parametrizations of cosmological quantities is in the fact that we are not comparing cosmological models within a certain framework of a parametric form and finding the best fit point in the parameter space and defining the confidence limits, but instead we are comparing an assumed cosmological model with its own possible variations. If the data suggest that the deviations from the actual model must be significant we can rule out the model right away. In fact parameterizations of cosmological quantities can never be used to falsify the assumed parametric form and we can only find the best fit point in the parameter space while the Bayesian interpretation of Crossing statistic is designed to falsify any assumed theoretical model. Crossing function can have different forms but considering the shape of the data and our expectations from the theoretical models (once again, $\mu(z)$ should increase smoothly and monotonically by redshift) we expect a simple Chebyshev polynomials [22] can perform satisfactorily. Chebyshev polynomials have properties of orthogonality and convergence within the limited range of $-1 < x < 1$ which is appropriate in our case of study. One can use Gram-Schmidt process to generate a set of orthogonal basis tailored appropriately for the case of study but in our analysis and considering the range of the data Chebyshev polynomials are good enough.

To falsify a theoretical model one needs to fit the given theoretical model along with the Crossing functions of different degrees. Looking at the data having more than two crossing between the actual model and the assumed model seems to be

unrealistic [1] and even if it is possible for some special models, it would not be detectable by the data since there are uncertainties in the data and if two models be so close to each other, data cannot discriminate between them. However, one needs not to be worry about this since assuming a Crossing function of high degree increases the allowed $\Delta\chi^2$ with respect to the χ^2 of the best fit point for different confidence levels so automatically analysis will show that by increasing the degrees of freedom of the function form the data would not be sensitive to pick up the tiny differences.

In this analysis I assume Chebyshev polynomials of orders one and two as the Crossing functions and to show the performance of the method, I do test the spatially flat Λ CDM model for three sets of simulated data according to three dark energy models (all spatially flat universes), an evolving dark energy model or Kink model with the same parameters used in [23] ¹ with $\Omega_m = 0.27$, a model of dark energy with constant equation of state of $w(z) = 0.9$ with $\Omega_m = 0.27$ and finally Λ CDM model with $w(z) = -1$ and $\Omega_m = 0.27$ (correct model). Simulated data is based on future JDEM data with 2298 data points in the range of $0.015 < z < 1.7$ [24]. The Kink model I used in this analysis has a special form of the equation of state of dark energy that at low redshifts it converges to $w(z) = -1$ and at higher redshifts it smoothly changes to $w = -0.5$. This evolving equation of state of dark energy is not within the possibilities of CPL [25] or many other dark energy parametrizations. Hence using these parameterizations results to wrong reconstruction of dark energy if the data is based on the Kink model (look at upper-right panel of figure 6 in [23]).

Crossing functions are given by:

$$T_I(C_1, z) = 1 + C_1\left(\frac{z}{z_{max}}\right) \quad (2.2)$$

and

$$T_{II}(C_1, C_2, z) = 1 + C_1\left(\frac{z}{z_{max}}\right) + C_2\left[2\left(\frac{z}{z_{max}}\right)^2 - 1\right], \quad (2.3)$$

where z_{max} is the maximum redshift in the data set. The functions that we fit to the data are $\mu_{DE}^{T_I}(z) = \mu_{DE}(\Omega_{0m}, z) \times T_I(C_1, z)$ and $\mu_{DE}^{T_{II}}(z) = \mu_{DE}(\Omega_{0m}, z) \times T_{II}(C_1, C_2, z)$ and variables are Ω_{0m}, C_1 in case of one crossing and Ω_{0m}, C_1, C_2 in case of two crossings. $\mu_{DE}(\Omega_{0m}, z)$ represents the luminosity distance of the assumed dark

¹equation of state of dark energy in this particular model is given by:

$$w(z) = w_0 + (w_m - w_0) \frac{1 + \exp(\Delta_t^{-1}(1 + z_t)^{-1})}{1 - \exp(\Delta_t^{-1})} \times \left[1 - \frac{\exp(\Delta_t^{-1}) + \exp(\Delta_t^{-1}(1 + z_t)^{-1})}{\exp(\Delta_t^{-1}(1 + z)^{-1}) + \exp(\Delta_t^{-1}(1 + z_t)^{-1})} \right], \quad (2.1)$$

with the constants having the values $w_0 = -1.0$, $w_m = -0.5$, $z_t = 0.5$, $\Delta_t = 0.05$.

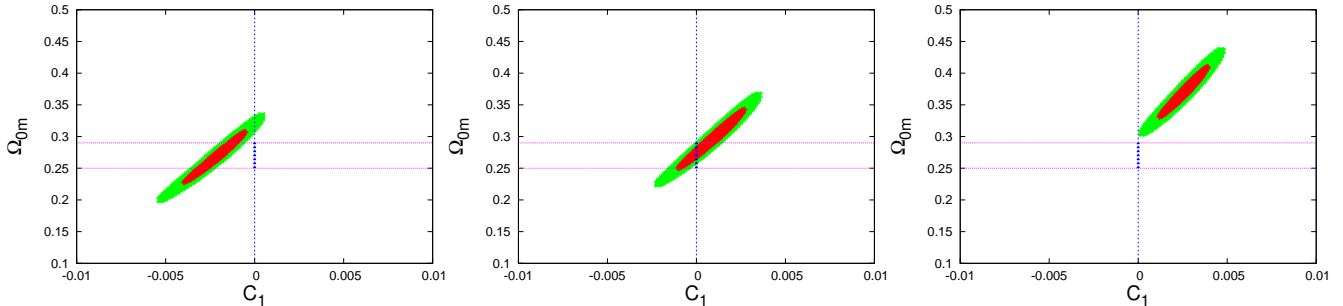


Figure 1: 1σ and 2σ confidence limits of fitting $\mu_{DE}^{T_I}(z)$ assuming the theoretical model being Λ CDM to three different simulated data sets. In the left panel the data is generated based on kink dark energy model with $\Omega_{0m} = 0.27$, the data in the middle panel is based on the Λ CDM model ($w = -1$) with $\Omega_{0m} = 0.27$ and in the right panel data is based on quiescence dark energy model with $w = -0.9$ and $\Omega_{0m} = 0.27$. The vertical line at $C_1 = 0$ represents the consistency of the assumed model and the data. Two horizontal lines at $\Omega_{0m} = 0.25$ and $\Omega_{0m} = 0.29$ represents possible constraints we may be able to put on the matter density from other cosmological observations to help us falsifying cosmological models.

energy model with matter density of Ω_{0m} . As it was mentioned earlier one can go to higher orders of crossing but we will see that the simulated data we are using are not sensitive to the higher orders. The rest is the usual likelihood analysis based on χ^2 statistics. In [1] we have explored that χ^2 statistic is in fact equivalent to the last mode of Crossing statistic and in the current analysis in fact we use an approximation to transfer the information of the first and second Crossing modes via the Crossing function to the last mode which is χ^2 . One may argue that the form of Crossing function may not be appropriate (to map the wrong model to the actual model) but considering the form and range of the data Chebyshev polynomials seems to be the most appropriate choice to do this mapping to the level of indistinguishability between the models.

In figure.1 we see the results of fitting $\mu_{DE}^{T_I}(z)$ assuming the theoretical model being Λ CDM to three different simulated data sets. The data in the left panel is based on kink dark energy model with $\Omega_{0m} = 0.27$, the data in the middle panel is generated based on Λ CDM model ($w = -1$) with $\Omega_{0m} = 0.27$ and the right panel is for the data based on quiescence dark energy model with $w = -0.9$ and $\Omega_{0m} = 0.27$. All models are spatially flat and the data is according to the future JDEM data.

We can Clearly see that $w = -0.9$ model can be ruled out with high confidence and the kink dark energy model shows inconsistency with the data at one to two sigma with no knowledge of matter density. Having a knowledge of matter density from other cosmological observations we can rule out the kink dark energy model too as it suggest the value of matter density should be about 0.32 for zero value of C_1 for standard flat Λ CDM model.

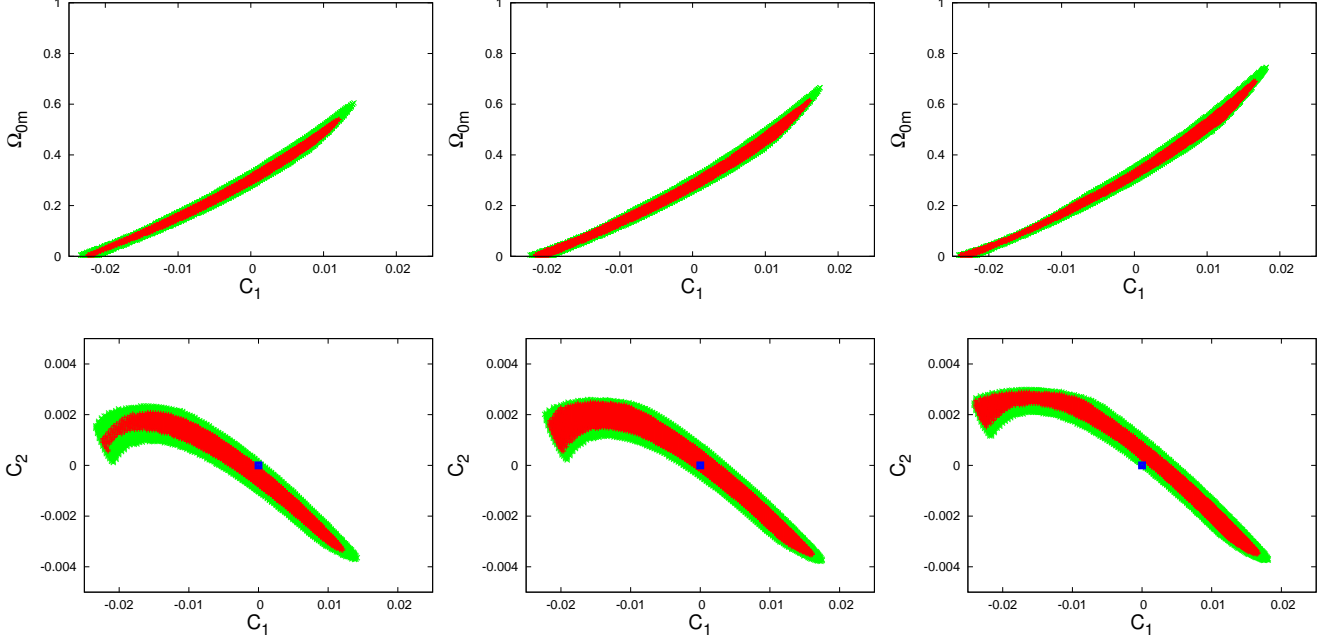


Figure 2: 1σ and 2σ confidence limits of C_1 - Ω_{0m} (upper panels) and C_1 - C_2 (lower panels) of fitting $\mu_{DE}^{T_{II}}(z)$ assuming the theoretical model being Λ CDM to three different simulated data sets similar to figure 1 (data is based on kink dark energy model in the left, Λ CDM in the middle and $w = -0.9$ in the right, all with $\Omega_{0m} = 0.27$). The blue dots in the lower panels represents $C_1 = C_2 = 0$ which shows the consistency of the assumed model and the data. One can see that the data are not sensitive to the second order function form T_{II} to distinguish between the cosmological models.

Assumed Model: Λ CDM			
	$\chi^2(\mu_{DE})$	$\chi^2(\mu_{DE}^{T_I})$	$\chi^2(\mu_{DE}^{T_{II}})$
Data based on Λ CDM Model with $\Omega_{0m} = 0.27$	2368.9	2368.2	2367.9
Data based on <i>Kink</i> Model with $\Omega_{0m} = 0.27$	2372.5	2368.7	2368.5
Data based on $w = -0.9$ Model with $\Omega_{0m} = 0.27$	2375.5	2368.9	2367.9

Table 1: Derived χ^2 using μ_{DE} , $\mu_{DE}^{T_I}$ and $\mu_{DE}^{T_{II}}$ fitting the three different simulated data sets. Note the changes in the χ^2 by increasing the degrees of freedom of the functions. Its clear that in all three cases, $\chi^2(\mu_{DE}^{T_I})$ and $\chi^2(\mu_{DE}^{T_{II}})$ are so close to each other that we can conclude the data is not sensitive to the second order of functions.

In figure 2 we see the results of fitting $\mu_{DE}^{T_{II}}(z)$ assuming the theoretical model being Λ CDM to three different simulated data sets similar to figure 1.

We see that by increasing the degrees of freedom of the Crossing function, Λ CDM model becoming more consistent to all three sets of the data which shows that the data is not sensitive to the crossing function of second order for these particular models of dark energy. I should emphasize that all of these three assumed dark

energy models are very close to each other when we allow matter density to vary [26]. For instance Λ CDM model with $\Omega_{0m} = 0.27$ is extremely close to kink dark energy model with $\Omega_{0m} = 0.314$. Looking at figure 2 also make it clear that there is no need to go to the higher orders of crossing. It basically means that the data cannot pick up two or three crossings of the comparing models that is clearly understandable and expected considering the quality and the range of the data.

To test the sensitivity of the method to the unknown dispersion of the data, σ_{int} , we re-do the analysis by increasing and decreasing the intrinsic dispersion of the data σ_{int} by 20% for all data points. In figure 3 we see the results for fitting $\mu_{DE}^{T_I}(z)$ based on spatially flat Λ CDM model to the data with under and over estimated sizes of the error-bars. As we see, the method still works robustly and can differ between the models accurately even when we have large uncertainties in estimation of the intrinsic dispersion of the data.

$w = -0.9$ model can be ruled out with high confidence and the kink model still shows inconsistency with the data at one to two sigma and having a knowledge of matter density from other cosmological observations we can rule out the kink model with high confidence.

3. Conclusion

In summary, we have presented a new approach based on Crossing Statistics that can be used to distinguish between different cosmological models using Crossing functions and hyper-parameters. We propose that instead of parametrizing the cosmological quantities such as luminosity distance, Hubble parameter or equation of state of dark energy, we can parametrize the possible deviations from any assumed model and compare a model with its own variations. Performance of the method is shown being very much promising in distinguishing cosmological models with and without knowledge of matter density in comparison with other parametric and non parametric methods. Using this approach one can falsify the standard model (or any particular dark energy model) given the data without knowledge of the underlying actual model of the universe and its parameters.

One may notice an interesting analogy between the first and second orders of Crossing functions in testing cosmological models using the cosmic distances and the spectral index and running of the spectral index of the primordial spectrum in CMB analysis. Similar to CMB analysis that we use the spectral index n_s to judge about the consistency of the scale invariant power spectrum (Harrison-Zeldovich form of the primordial spectrum) to the data by looking at the derived n_s and its error-bars (whether $n_s = 1$ is in agreement to the CMB data), we use C_1 to test the consistency of the assumed dark energy model to the data (whether $C_1 = 0$ is in agreement to the distance data). Similarly there is an analogy between the running of the spectral index in CMB studies and C_2 in our analysis. One may argue that n_s is a

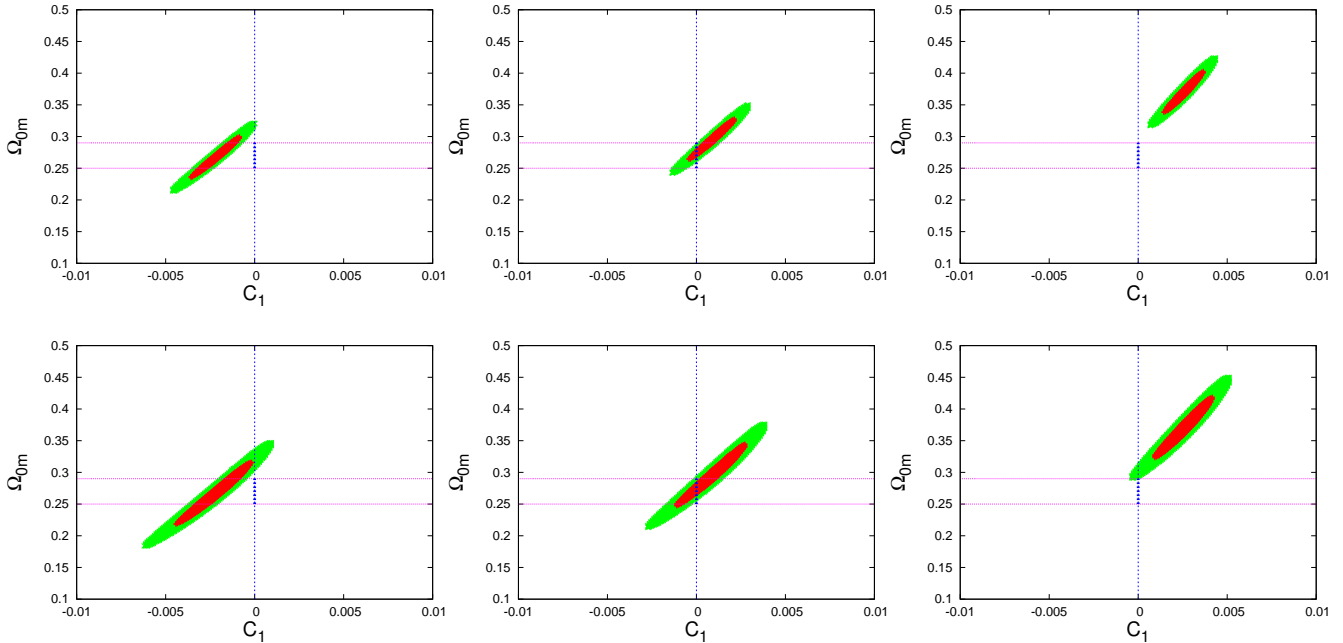


Figure 3: 1σ and 2σ confidence limits of fitting $\mu_{DE}^{T_I}(z)$ assuming the theoretical model being Λ CDM to three different simulated data sets similar to figure 1 where in upper panels we have reduced the size of the error-bars for all data points by 20% and in the lower panels we have increased the size of the error-bars for all data points by 20%. We can see that the method is robust against the uncertainties in the data up to a very high level.

physical parameter while C_1 is not. We should note that independent of considering inflationary scenarios (or any physical mechanism resulting to a power-law form of the primordial spectrum) one can still use n_s from a phenomenological perspective to rule out Harrison-Zeldovich form of the primordial spectrum.

The Bayesian interpretation of the Crossing Statistic appears to us to be a promising method of confronting cosmological models with supernovae observations, and has the potential to be trivially generalized to other cosmological observations.

Acknowledgments

AS thanks E. Linder, D. Pogosyan, L. Perivolaropoulos, T. Clifton, P. Ferreira and U. Seljak for useful discussions. This work has been supported by World Class University grant R32-2009-000-10130-0 through the National Research Foundation, Ministry of Education, Science and Technology of Korea.

References

- [1] A. Shafieloo, T. Clifton and P. Ferreira, JCAP **08**, 017 (2011).

- [2] V. Sahni and A. A. Starobinsky, Int. J. Mod. Phys. D **15**, 2105 (2006).
- [3] R. A. Daly and S. G. Djorgovski, Astrophys. J. **597**, 9 (2003).
- [4] D. Huterer and G. Starkman, Phys. Rev. Lett. **90** 031301 (2003).
- [5] Y. Wang and P. Mukherjee, Astrophys. J. **606**, 654 (2004).
- [6] Y. Wang and M. Tegmark, Phys. Rev. D **71**, 103513 (2005).
- [7] A. Shafieloo, U. Alam, V. Sahni and A. A. Starobinsky, Mon. Not. Roy. Astron. Soc. **366**, 1081 (2006).
- [8] A. Shafieloo, Mon. Not. Roy. Astron. Soc. **380**, 1573 (2007).
- [9] V. Sahni, A. Shafieloo and A. A. Starobinsky, Phys. Rev. D **78**, 103502 (2008).
- [10] C. Zunckel and C. Clarkson, Phys. Rev. Lett **101**, 181301 (2008).
- [11] A. Shafieloo, V. Sahni and A. A. Starobinsky, Phys. Rev. D **80** R, 101301 (2009).
- [12] C. Clarkson and C. Zunckel, Phys. Rev. Lett. **104**, 211301 (2010).
- [13] A. Shafieloo and C. Clarkson, Phys. Rev. D **81**, 083537 (2010).
- [14] S. Nesseris and A. Shafieloo, Mon. Not. Roy. Astron. Soc. **408**, 1879 (2010).
- [15] T. Holsclaw *et. al*, Phys.Rev. Lett **105**, 241302 (2010).
- [16] R. Crittenden *et al*, arXiv:1112.1693 (2011).
- [17] T. M. Davis, Astrophys. J. **666**, 716 (2007).
- [18] P. Serra, A. Heavens and A. Melchiorri, Mon. Not. Roy. Astron. Soc. **379**, 169 (2007).
- [19] E. Linder and R. Miquel, Int. J. Mod. Phys. D **17**, 2315 (2008).
- [20] J. Sollerman, Astrophys. J. **703** 1374 (2009).
- [21] M. Kilbinger *et al*. Mon. Not. Roy. Astron. Soc. **405**, 2381 (2010).
- [22] R. J. Barlow, *Statistics*, John Wiley & Sons Ltd, (1989).
- [23] T. Holsclaw *et. al*, Phys. Rev. D **82**, 103502 (2010).
- [24] <http://jdem.lbl.gov/>
- [25] M. Chevallier and D. Polarski, Int. J. Mod. Phys. D **10**, 213 (2010); E.V. Linder, Phys. Rev. Lett. **90**, 091301 (2003).
- [26] A. Shafieloo and E. Linder, Phys. Rev. D **84**, 063519 (2011).